

# Découverte et exploitation de la hiérarchie des tâches par motivation intrinsèque

Nicolas Duminy<sup>1,2</sup> and Sao Mai Nguyen<sup>2,3,4</sup>

<sup>1</sup>Université Bretagne Sud, Lorient, France.

<sup>2</sup>Lab-STICC, UMR 6285

<sup>3</sup>IMT Atlantique, Brest, France ,

<sup>4</sup>U2IS, ENSTA, IP Paris & Inria, FLOWERS team,  
nguyensmai@gmail.com

Juin 2020

Les tâches des robots peuvent être de types variés, hiérarchiques, et peuvent subir des changements radicaux ou même être créées après le déploiement du robot. Le robot doit faire face à des tâches avec des complexités variables, certaines demandant une action simple pour être réalisée, d'autre demandant une séquence d'actions. Une méthode permettant de guider ce choix se nomme la motivation intrinsèque. Le robot est guidé vers les zones les plus intéressantes de son environnement afin d'apprendre les compétences les plus intéressantes. Il est capable d'évaluer la complexité de l'action nécessaire pour réaliser une tâche. Quand il fait face à des tâches hiérarchiques de différentes complexités, qui peuvent être réalisées par une combinaison de tâches plus simples, le robot utilise une nouvelle manière d'acquérir des compétences en explorant la hiérarchie des tâches elle-même, en combinant ses compétences via une combinaison de tâches afin d'acquérir des nouvelles et plus complexes.

Pour l'apprentissage de séquences d'actions pour réaliser de multiples tâches de complexités différentes, en apprenant et exploitant la hiérarchie des tâches grâce à la nouvelle infrastructure algorithmique que nous introduisons : les procédures, qui permet la combinaison de séquences d'actions connues en fonction des effets de ces actions. Ceci mène au développement de deux algorithmes [1]. Le premier, appelé Intrinsically Motivated Procedure Babbling (IM-PB), permet de montrer que hiérarchie des tâches peut être apprise de manière autonome en même temps que l'espace des actions motrices. Le second, appelé Socially Guided Intrinsic Motivation with Procedure babbling (SGIM-PB), montre que cette exploration autonome de la hiérarchie des tâches et de l'espace des actions complexes peut être accéléré par des experts humains fournissant des démonstrations. Ces deux implémentations sont testés sur un bras à 3 degrés de liberté en simulation pour réaliser 6 types de tâches différentes : contrôle

de la position 3D de l'effecteur, d'un stylo, dessiner, contrôler des joystick et contrôler un personnage de jeux vidéo.

L'algorithme SGIM-PB est également testé sur le robot industriel Yumi apprenant des tâches hiérarchiques avec des séquences d'actions motrices dans un environnement physique: contrôle de la position de l'effecteur, déplacement d'objets, produire des sons. Ce test fut d'abord réalisé en simulation, puis confirmé sur le robot réel. Une vidéo est disponible sur <https://www.youtube.com/watch?v=29KDTeAkZis>

Ainsi, l'architecture proposée est capable d'apprendre des compétences en reliant les mouvements réalisés, appelés actions ou politiques, aux conséquences observées sur son environnement. Ce travail contribue aux domaines de :

- Apprentissage de séquences d'actions motrices : nous proposons une infrastructure algorithmique appelée *procédure*, développée pour permettre à un apprenant stratégique et interactif de découvrir et exploiter la hiérarchie des tâches, en combinant des actions connues en fonction des tâches. Nous lui avons également permis de construire des séquences d'actions de n'importe quelle taille, lui permettant d'adapter cette taille à la tâche en cours d'étude;
- Apprentissage hiérarchique : nous avons proposé une nouvelle représentation de la hiérarchie des tâches avec les procédures. Nous avons montré que SGIM-PB et IM-PB peuvent apprendre cette hiérarchie, et qu'elle permet un transfert de connaissance des tâches simples vers les tâches complexes.
- Apprentissage interactif : nous avons développé une nouvelle manière de fournir des démonstrations à un robot pour un expert humain, en utilisant la hiérarchie des tâches via les procédures. Nous avons montré dans [2] que les démonstrations les plus utiles par les experts humains sont les actions bas niveau dans le cas de tâches simples, et sont la décomposition de tâches dans le cas de tâches complexes et hiérarchique;
- Apprentissage de curriculum : nous montrons que SGIM-PB et IM-PB peuvent déterminer par eux-mêmes d'apprendre les tâches simples avant les tâches complexes, mais aussi de choisir s'il peut réutiliser les connaissances apprises pour les tâches complexes

## References

- [1] Nicolas Duminy, Sao Mai Nguyen, and Dominique Duhaut. "Effects of social guidance on a robot learning sequences of policies in hierarchical learning". In: *International Conference on Systems Man and Cybernetics*. Ed. by IEEE. 2018.
- [2] Nicolas Duminy, Sao Mai Nguyen, and Dominique Duhaut. "Learning a Set of Interrelated Tasks by Using a Succession of Motor Policies for a Socially Guided Intrinsically Motivated Learner". In: *Frontiers in Neurorobotics* 12 (2019), p. 87. ISSN: 1662-5218. DOI: 10.3389/fnbot.2018.00087. URL: <https://www.frontiersin.org/article/10.3389/fnbot.2018.00087>.